INTELLIGENT BUDDY

Kevin Roy
Department of Computer Science and Engineering
Amal Jyothi College of Engineering
Kanjirappaly, Kerala
kevinroy2023@cs.ajce.in

Lino Shaji
Department of Computer Science and Engineering
Amal Jyothi College of Engineering
Kanjirappaly, Kerala
linoshaji2023@cs.ajce.in

Riya G Johnson
Department of Computer Science and Engineering
Amal Jyothi College of Engineering
Kanjirappaly, Kerala
riyagjohnson2023@cs.ajce.in

Tince Tomy
Department of Computer Science and Engineering
Amal Jyothi College of Engineering
Kanjirappaly, Kerala
tincetomy2023@cs.ajce.in

Jane George
Department of Computer Science and Engineering
Amal Jyothi College of Engineering
Kanjirappaly, Kerala
janegeorge@amaljyothi.ac.in

***Abstract*—Software agents known as voice assistants are able to understand human speech and answer with synthetic voices. The most well-known voice assistants, which are built into smartphones or specific home speakers, are Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant. Users can use voice commands to manage other basic chores like email, to-do lists, and calendars, as well as ask inquiries of personal assistants, operate home automation devices, and playback of media. By holding and analysing information in the context of the user, this engages the capacity for social communication through natural language processing. This research will examine the fundamental operations and typical characteristics of voice assistants in use today. The currently used system operates online and is kept up by a third party. This program will safeguard personal information from others and use the local database, speech recognition and synthesizer.**

***Index Terms*—Speech recognition system, Voice assistants, Machine learning**

I. Introduction

A voice assistant is a type of digital assistant that listens for specific voice commands, language processing algorithms, and synthesis to find relevant information or carry out specific tasks as requested by the user. Supported commands, also referred to as "intents", are told by the user, and voice assistants find relevant information by hearing for specific keywords and removing background noise. While voice assistants may be entirely a software system on which all devices can be combined, other assistants are specifically designed for each individual device use, such as the Amazon Alexa. Voice assistants are now incorporated into a number of the gadgets we use on a daily basis, including smartphones, PCs, and high-quality speakers. In this work, we hope to create a speech recognition-based assistant that can help users with everyday activities and introduce them to online resources for learning.

II. Literature Survey

A Smart Personal Assistant by Shrooti Singh, is a digital Assistant for visually faulty human beings in an indoor environment. The voice assistant system is built with the Python programming language and implemented with the PyCharm IDLE [1].The paper discusses a smart personal assistant designed for visually impaired individuals in an indoor setting. The system is built using Python and PyCharm IDLE and uses a well-designed NLU algorithm with deep learning to interpret ambiguous responses from users, allowing the assistant to provide accurate answers.

Desktop Voice Assistant for Visually Impaired [2] by Ankur Sindhu, uses google API to convert input speech into text. The paper describes a desktop voice assistant designed for visually impaired users, which uses Google API to convert speech to text. The system sends the converted text to Google Cloud for processing and receives the resulting output. The system then identifies whether the command is for the system or a web browser and sends the output to the Python backend, which provides the desired output to the user.

Voice Assistant using Artificial Intelligence [3] by A

ullas, gathers the audio from the microphone and then convert that into text, later it is sent through GTTS (Google text to speech). The paper describes a voice assistant system that uses artificial intelligence. The system records audio from the microphone, converts it into text, and then uses Google Text-to-Speech to convert the text into an audio file in English language. The audio is then played using the play sound package in Python. AI-based voice assistants are operating systems that can recognize human voice and respond using integrated voices.

SPYDER: Intelligent Voice Assistant by Piyush Gupta [4], uses Python libraries and Speech Recognition APIs are used to integrate the personal voice assistant. The paper describes SPYDER, an intelligent voice assistant that integrates Python libraries and Speech Recognition APIs. Python Text to Speech module is used to translate voice responses, and the system's Graphical User Interface (GUI) is created using the QT designer tool, which is an open-source GUI builder that generates UI files rather than code in any programming language. The proposed system aims to provide a personal voice assistant to users.

Life Assistant for Visually Impaired People using AI [5] by Bhanushali V, uses Natural Language Understanding(NLU) and Voice Recognition. Built-in falling detection algorithm based on gyroscope.The paper describes a life assistant designed for visually impaired individuals that uses Natural Language Understanding (NLU) and Voice Recognition. The system also includes a built-in falling detection algorithm based on gyroscope data. The proposed system aims to assist visually impaired individuals in their daily lives by providing them with an intelligent assistant that can understand and respond to their commands and detect falls to ensure their safety.

A vision and speech enabled, customizable, virtual assistant for smart environments [6] by Mario Grasso, graph shows expressions and is enabled with speech synthesis and recognition, face detection and face recognition for user identification.The paper describes a vision and speech-enabled virtual assistant designed for smart environments, which is customizable to meet user needs. The system features a graph that shows expressions and is enabled with speech synthesis and recognition, face detection, and face recognition for user identification. The proposed system aims to provide an intelligent assistant that can interact with users and assist them in various tasks in smart environments.

Speech based virtual assistant system for visually impared people [9] by P Sumalatha, uses speech based virtual assistant for helping and guiding the visually impaired people for their daily tasks when required. The paper describes a speech-based virtual assistant system designed to assist and guide visually impaired individuals in their daily tasks. The system presents a modular solution for improving web-based accessibility for visually impaired persons. It uses text-to-speech and speech-to-text interfaces to allow users to communicate and customize with the system. The system includes five modules, namely Voice-based chatbot, face detection, object recognition, text recognition, and website access, that are currently implemented. The proposed system aims to provide visually impaired individuals with a user-friendly and accessible interface to interact with and access web-based content.

Spectral Voice Conversion For Text to Speech Synthesis [10] by A Kain, uses new spectral conversion algorithm using a locally linear transformation based on Gaussian mixture models whose parameters are trained by joint density estimation. The paper presents a new spectral voice conversion algorithm for text-to-speech synthesis using a locally linear transformation based on Gaussian mixture models. The model parameters are trained by joint density estimation. Numerical experiments showed that the proposed algorithm performed similarly to a previous GMM-based approach but was more robust for small amounts of training data. The proposed algorithm aims to improve the quality of text-to-speech synthesis by using a more robust and efficient method for spectral voice conversion.

III. METHODOLOGY

The Speech Recognition library has many built-in features that will enable the assistant to understand the command given by the user and the response will be sent back to the user in voice, using Text to Speech functions, in the proposed concept effective way of implementing a Personal voice assistant. The algorithms used in the background will translate the user's spoken instruction into text when the assistant hears it. The following features are intended for the proposed system: 1) Using the microphone to capture speech patterns. 2) conversion from audio data recognition. 3) Assessing the input against pre-established commands. 4) Producing the required results. The overall architecture of the system is illustrated in Figure 3.1.

1) ***Taking the input as speech patterns through microphone****:* While continuing to listen for commands, the assistant solicits input from the user. The amount of time for listening can be adjusted based on the needs of the user. If the assistant doesn't understand the command correctly, it will keep asking the user to repeat it. Depending on the user's preferences, the voice of this assistant can be changed to be either male or female. To enable the use of speech synthesis and recognition within Windows applications, Microsoft created the Speech Application Programming Interface, or SAPI.

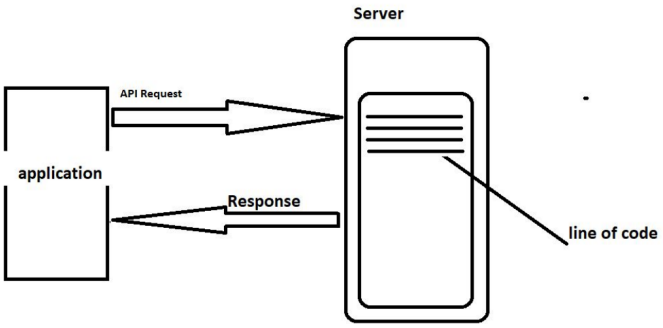

Figure 3.1 Architecture of the system

*2) **Audio data recognition and conversion into text**:* Speech recognition, also known as automatic speech recognition (ASR), computer voice recognition, or speech-to-text, is the ability of a program to translate spoken words into written language. Despite the fact that the two terms are occasionally used interchangeably, voice recognition only seeks to identify a specific person's voice, whereas speech recognition focuses on transforming speech from a verbal to a text format. In our project, we'll be using AssemblyAI to recognize speech. The quality and accessibility of the lauded AssemblyAI voice recognition API are unmatched. The AssemblyAI API can be used to understand and transcribe audio and video files using AI models. Since it is free and easy to use, this API was chosen.

*3) **Comparing the input with predefined commands**:* Matching spoken words or phrases to a list of predetermined commands is how input is compared to established commands. Natural methods are frequently used for this approaches for natural language processing (NLP), including speech recognition and intent categorization. The system analyzes user input using algorithms, finds the appropriate command, and executes that command. This method enables the voice assistant to comprehend the user's intent and react suitably, giving it a practical and user-friendly tool for carrying out a variety of tasks.

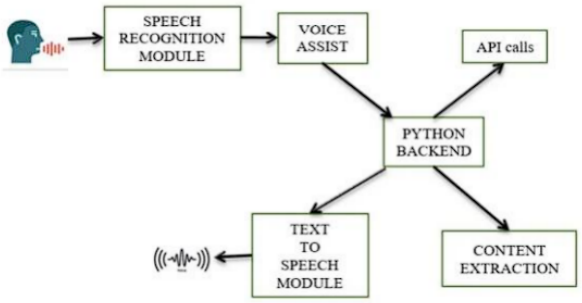

Figure 3.2: Detailed Workflow of the voice assistant

*4) **Giving the desired output**:* The voice recognition module's output is obtained by the Python backend, which then determines if the command or the speech output is an API call and context extraction. The output is then returned to the Python backend so that the user can receive the desired output. The desired output of a voice assistant is produced utilizing Text-to-Speech (TTS) technology. A language model like myself texts an answer, which the voice assistant then uses TTS software to turn into speech. Based on the supplied text, the TTS software creates human-like speech that is then output as audio through the device's speaker. Instead of needing to read the voice assistant's comments and information, users may converse with it and hear what it has to say.

IV. ARCHITECTURE

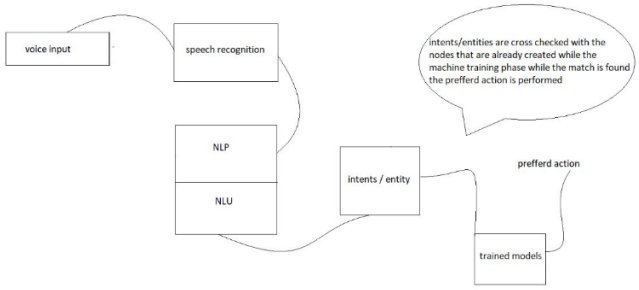

Figure 3.3 Detailed Architecture of the system

After identifying the intents and entities in natural language input, the next step is to perform the intended activity. This is achieved by comparing the identified intents/entities with predefined or user-defined functions. For example, a user-defined function called ceate Todo()" may contain a program that performs the action of creating a task, and this function may have several user-defined nodes with functions for performing different activities. These functions are added during the machine training phase when the machine encounters commands that are not defined. The machine notifies the admin, who can update the database with the required information. The front-end of a Virtual Personal Assistant (VPA) that interacts with users must be able to comprehend human language by converting the given natural language input into relevant representations. Natural language processing (NLP) techniques such as tokenization, stemming, and part-of-speech tagging can be used to convert the input into a structured format that can be understood by the machine. Machine learning algorithms such as decision trees and neural networks can then be trained on this structured data to identify intents and entities accurately.

Once the intents and entities have been identified, the machine can retrieve relevant information from the database and perform the intended activity. For example, if the user asks the VPA to create a task, the machine can retrieve the necessary information such as the task name, due date, and priority level from the database and create the task accordingly. Overall, the successful operation of a VPA depends on its ability to accurately comprehend and process human language, and this is achieved through a combination of NLP techniques and machine learning algorithms.

V. Experimental results

Fig. 1. Main User Interface of the app

The above figure shows the user interface of the application

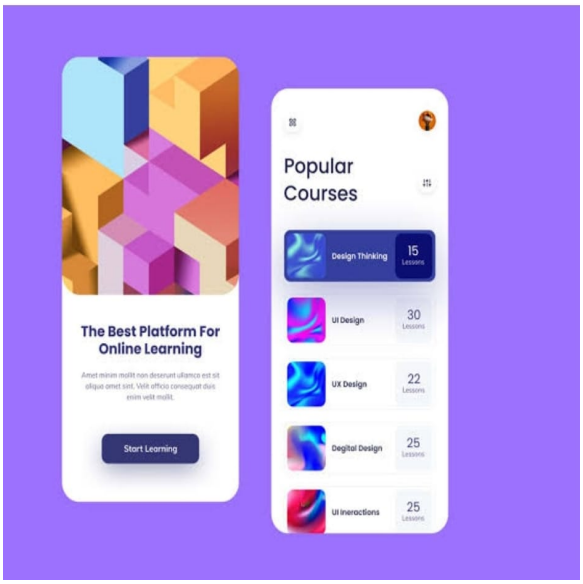

Fig. 2. Menu of the application

fig 2 shows the menu of the application for the selection of courses for the students

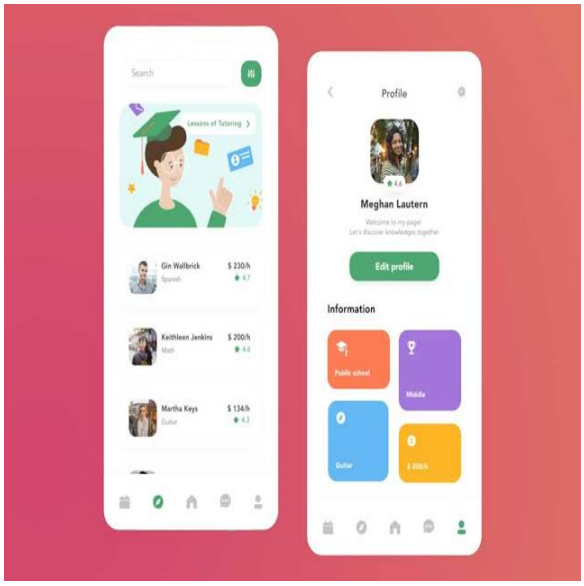

Fig. 3. Dashboard of the application

fig 3. shows the information about the students using the application

VI. Conclusion

Python-based personal assistant work discussed in this paper provides an efficient and supportive solution for saving time and assisting individuals with disabilities. By leveraging voice-activated technology, the personal assistant can complete assigned tasks quickly and easily with just a voice command. The assistant is capable of performing a wide range of tasks, including sending messages, automating YouTube, and retrieving information from Google and Wikipedia. This automation has made tasks such as web searching and weather forecast retrieval much easier for users. Moving forward, the aim of this project is to continue improving the performance of the personal assistant by adding more functions and methods, making it even more efficient, reliable, and capable of dealing with various situations. With ongoing development and advancements in technology, voice-activated personal assistants are poised to become even more widespread and indispensable in our daily lives.

References

[1] **Saunshimath, Nirmala Thakur, Sakshi Singh, Shrooti Singh, Ranavijay .K, Kumuda.** (2022). SAHARA -A Smart Personal Assistant. 10.13140/RG.2.2.33808.48642.

[2] **Yadav, A., Singh, A., Sharma, A., Sindhu, A., Rastogi, U. (2020).** Desktop voice assistant for visually impaired. Int J Recent Technol Eng (IJRTE), 9(2)..

[3] **G. Preethi , Abishek. K , Thiruppugal S , Vishwaa D A, 2022**, Voice Assistant using Artificial Intelligence, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) Volume 11, Issue 05 (May 2022)

[4] **Piyush Gupta, Sakshi Yadav** "SPYDER: Intelligent Voice Assistant," International Research Journal of Engineering and Technology (IRJET) Volume 08 Issue 04 (April 2021)

[5] **Bhanushali Nayan, V., Kapile Namrata, R., Pawar Shubham, G., Pawase Kalpesh, D., Tambe Prachi**. LIFE ASSISTANT FOR VISUALLY IMPAIRED PEOPLE USING AI.

[6] **Iannizzotto, G., Bello, L. L., Nucita, A., Grasso, G. M. (2018, July)**. A vision and speech enabled, customizable, virtual assistant for smart environments. In 2018 11th International Conference on Human System Interaction (HSI) (pp. 50-56). IEEE.

[7] **Safnaz, Nayana Manju Jogi, Shreya Shetty, Prasad Rathode, Deeksha K R, 2021**, Blind Assistance Device using AI, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) NCCDS – 2021 (Volume 09 – Issue 12),

[8] **Chen, R., Tian, Z., Liu, H., Zhao, F., Zhang, S., Liu, H. (2018, May).** Construction of a voice driven life assistant system for visually impaired people. In 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 87-92). IEEE.

[9] **S. Subhash, P. N. Srivatsa, S. Siddesh, A. Ullas and B. Santhosh,** "Artificial Intelligence-based Voice Assistant," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 593-596, doi: 10.1109/WorldS450073.2020.9210344.

[10] **M. R. Sultan and M. M. Hoque,** "ABYS(Always By Your Side): A Virtual Assistant for Visually Impaired Persons," 2019 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038603